Interim report of the Recruitment Working Group

Grant Thompson

September 23, 2014

Introduction

- In 2012, the Groundfish Plan Teams and Crab Plan Team appointed a working group (Bob Foy, Jim Ianelli, Diana Stram, Grant Thompson) to list/evaluate alternatives for a number of assessment and management issues related to recruitment

- Phases I and II of the working group's report were structured around the list of items considered at an April 2012 workshop, and were submitted in May and September of 2012, respectively

- Other working group members (Anne Hollowed, André Punt, William Stockhausen, Farron Wallace) were added following completion of Phase II

- Phase III of the report was submitted in September 2013

- After reviewing the Phase III report, the Groundfish Plan Teams had suggestions for further work on items B1, B2, B4, B5, B7, and C1

- The Teams' suggestions for two of these items (B1 and B7) were addressed by the working group in 2014

- Because only two items are addressed, the present report is referred to as an "interim" report rather than a full Phase IV report

- Presentation format for each item: problem statement slide, followed by a history slide, followed by a set of slides emphasizing what is *new* in the interim report

NOAA FISHERIES

B1: Establishing criteria for excluding individual within-regime year classes from estimates

Problem statement

- Age-structured assessments produce a time series of estimated recruitments, but recent recruitments may be estimated poorly, and if they are used to estimate SR parameters or the parameters of a spawning-independent recruitment distribution without being weighted appropriately (e.g., as in Proj), biased estimates may result

Year	Mean	Variance	Inverse variance	Inverse-variance-weighted mean
2002	4.91	2.0	0.50	2.46
2003	5.04	2.1	0.48	2.40
2004	5.19	2.2	0.45	2.36
2005	5.32	2.3	0.43	2.31
2006	5.24	2.4	0.42	2.18
2007	5.29	2.5	0.40	2.12
2008	4.58	2.6	0.38	1.76
2009	5.02	2.7	0.37	1.86
2010	5.12	2.8	0.36	1.83
2011	4.80	2.9	0.34	1.66
2012	4.47	3.0	0.33	1.49
2013	20.00	100.0	0.01	0.20
Mean:	6.25			5.05
Error:	0.25			0.01

- How many recent year classes should be excluded from the time series?

NOAA FISHERIES

History

- Phase I: Provisional recommendation was to exclude the 4 most recent years in the time series
- 6/12: SSC *"suggested that the Plan Teams consider life history when selecting the years to exclude from the time series"*
  - What if we do age-structured assessments of species with 1-year lifespans?
- Phase II: Provisional recommendation was switched to a formula based on the natural mortality rate (M) and the first age at which survey selectivity was at least 10% ($A_{10\%}$), with the idea that year classes with current age < $A_{10\%}$ should be excluded from the time series, at a minimum
  - The formula was loosely based on a few example stock assessments
- Phase III: "Provisional recommendation" was upgraded to "recommendation"
- 9/13: Groundfish Plan Teams *"recommended that the working group conduct some further analysis and that the working group consider another alternative which uses* $A_{50\%}$ *(age at 50% selectivity)"*
  - In other Team discussion (not in the minutes), interest was also expressed in basing the formula on a wider range of stock assessment examples
- This year: Several equations were tested against a wider range of stock assessment examples, using $A_{10\%}$ and $A_{50\%}$ as alternative reference ages

NOAA FISHERIES

Data

- In August 2014, assessment authors were asked to provide the following items of information for each of their age-structured assessments:
  - The natural mortality rate (M)
  - The first age at which survey selectivity reaches 10% ($A_{10\%}$)
  - The first age at which survey selectivity reaches 50% ($A_{50\%}$)
  - The first age for which the recruitment corresponding to the model's estimated abundance in the current year is included in the recruitment series ($first\_age$)
    - For example, if the assessment model begins with age 1, but the estimated recruitments corresponding to ages 1 and 2 in the current year are not included in the recruitment time series, then $first\_age$=3
- Authors responded by providing data for 26 assessments (15 BSAI, 10 GOA, 1 both)
  - In cases where separate values were provided for males and females, the values were averaged
  - In cases where survey selectivity at the first age in the assessment model was greater than 10% (or 50%), $A_{10\%}$ (or $A_{50\%}$) was set at the first age in the model

Factorial design

- In the overall data set, three factors emerged as being potentially important for individual consideration when analyzing the data:
  - In the 2013 assessment of GOA Pacific cod, the author changed $first\_age$ from the traditional value of 2 to 4, because she thought that she was required to do so by the Phase III report
  - The Phase III report used 10% as the minimum survey selectivity that should be accepted for estimation of incoming year classes, but the Plan Teams recommended adding a survey selectivity of 50% to the analysis
  - Many of the assessments use estimates of abundance at ages lower than either of the suggested cutoffs: 35% use model estimates of abundance at ages less than $A_{10\%}$, and 77% use model estimates of abundance at ages less than $A_{50\%}$
    - For $A_{10\%}$, excluding such data would reduce the sample size from 26 to 17, and for $A_{50\%}$, excluding such data would reduce the sample size from 26 to 6
- This resulted in a factorial design of 2^3=8 different ways of looking at the data:
  - GOA Pcod = {2,4}, selectivity cutoff = {10%,50%}, ages<cutoff = {include,exclude}

Models

- The following models were fit to the data ($A\%$ represents either $A_{10\%}$ or $A_{50\%}$, depending on which value is used in any given combination of the three factors):

$$Model\,1: \quad first\_age = \frac{a_1}{M + b_1}$$

$$Model\,2: \quad first\_age = \exp(a_2 \cdot A\% + b_2)$$

$$Model\,3: \quad first\_age = \left(\frac{a_3}{M + b_3}\right) \cdot \exp(c_3 \cdot A\% + d_3)$$

$$Model\,4: \quad first\_age = \frac{a_4 \cdot A\%}{(M + b_4) \cdot (A\% + c_4)}$$

$$Model\,5: \quad first\_age = a_5 + b_5 \cdot M + c_5 \cdot A\% + d_5 \cdot M \cdot A\%$$

$$Model\,6: \quad first\_age = a_6 + \frac{b_6}{M} + c_6 \cdot A\% + \frac{d_6}{M} \cdot A\%$$

$$Model\,7: \quad first\_age = \frac{a_7}{M} + A\%$$

Results (1 of 2)

- Three statistics relating to goodness of fit for each model within each combination of the three factors are shown in Table G.1 (models are sorted in decreasing order):
  - Akaike's Information Criterion (AIC)
  - The coefficient of determination (R^2)
  - "Error," which describes the proportion of the data where the predicted value of $first\_age$ is less than $A\%$
- Focusing on AIC, Model 7 ranked as the best model whenever data with $first\_age<A\%$ were excluded, Model 2 ranked as the best model when such data were not excluded and $A\%$ was set equal to $A_{10\%}$, and Model 4 ranked as the best Model when such data were not excluded and $A\%$ was set equal to $A_{50\%}$
- Focusing on R^2, Model 6 ranked as the best model for all combinations of factors except when $A\%$ was set equal to $A_{50\%}$ and data were excluded whenever $first\_age$ was less than $A\%$, in which cases Model 5 ranked as the best model
  - Model 7 ranked no lower than third when $first\_age<A\%$ data were excluded
- Focusing on "Error," Model 7 ranked as the best model (with Error=0) whenever data with $first\_age<A\%$ were excluded; whereas when data with $first\_age<A\%$ were not excluded, a variety of models ranked near the top, but it should be noted that the Error values for the top-ranked models in all such cases were very high

NOAA FISHERIES

Results (2 of 2)

- By all three criteria, then, Model 7 does very well if $first\_age<A\%$ data are excluded
- The fit of Model 7 to the data with GOA Pcod $first\_age$=2, $A\%=A_{10\%}$, and data with $first\_age<A\%$ excluded is shown below:
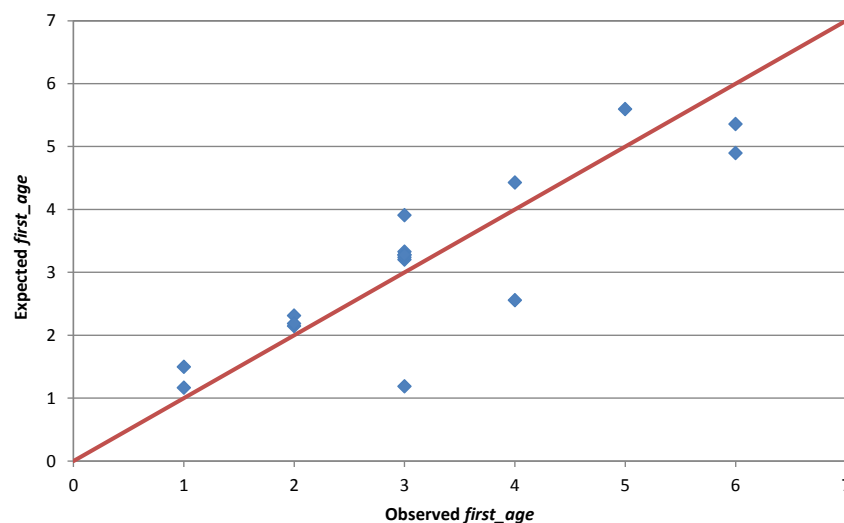
- With only one parameter, Model 7 is also easy to extend to the more realistic case where the right-hand side of the equation is constrained to take integer values
  - Rounding the values on the right-hand side to the nearest integer and profiling over a_7 (for the same combination of factors listed above) indicates that a_7=0.05 does as well as any other value of a_7 (Working Group recommendation)

B7: Preferred measure of central tendency in recruitment

Problem statement

- BSAI and GOA Groundfish FMPs specify that Tier 3 calculation of $B_{40\%}$ be based on mean recruitment

- Both Proj and Stock Synthesis also base reference points on mean recruitment

- For some stocks, use of median recruitment gives a *very* different estimate of stock status than the mean

  - E.g., AI pollock, BSAI Greenland turbot, AI blackspotted/rougheye rockfish

- Which estimator of central tendency is "better?"

History

- Phases I-II: Item B7 was not included, having not been on the agenda of the April 2012 workshop

- 11/12: BSAI Plan Team *"recommends that the Recruitment Working Group examine use of median recruitment (or other measure(s) of central tendency) as an alternative to mean recruitment for calculation of reference points"*

- Phase III: Item B7 was added in response to BSAI Plan Team recommendation
  - Simulation analysis (Appendix F) compared use of median and mean recruitment in terms of short-term and long-term mean and standard deviation of biomass, exploitation, and catch; and 95% CI, median, and mean of rebuilding time
    - True values of median and mean recruitment were assumed known
  - Working Group recommended staying with mean recruitment

- 9/13: Groundfish Plan Teams *"recommended further analyses, in particular expanding Appendix F to include real-time updating of estimates of the mean and median (rather than assuming that the true values are always known)"*

- This year: Appendix F was revised to include real-time updating, and a second analysis was added to illustrate some worst-case scenarios

NOAA FISHERIES

Analysis #1: recruitment distribution (1 of 2)

- To capture the idea of occasional recruitments that are much larger or more common than would be expected from a single lognormal distribution, recruitments were drawn randomly from a weighted sum of two lognormal distributions, one representing "typical" recruitments and the other representing "extreme" recruitments

- Parameter values governing the recruitment distribution were as follow:
  - For the "typical" distribution, μ = 0, σ = 0.6
  - For the "extreme" distribution, μ = 2, σ = 0.05
  - Proportion of time that recruitment is "typical" = 0.95

- The ratio of the median "extreme" recruitment to the median "typical" recruitment is about 7.4, and the ratio of the means is about 6.2

- The ratio of the *lower* end of the 95% confidence interval for the "extreme" distribution to the *upper* end of the 95% confidence interval for the "typical" distribution is about 2.1

- See figure on next slide

Analysis #1: recruitment distribution (2 of 2)
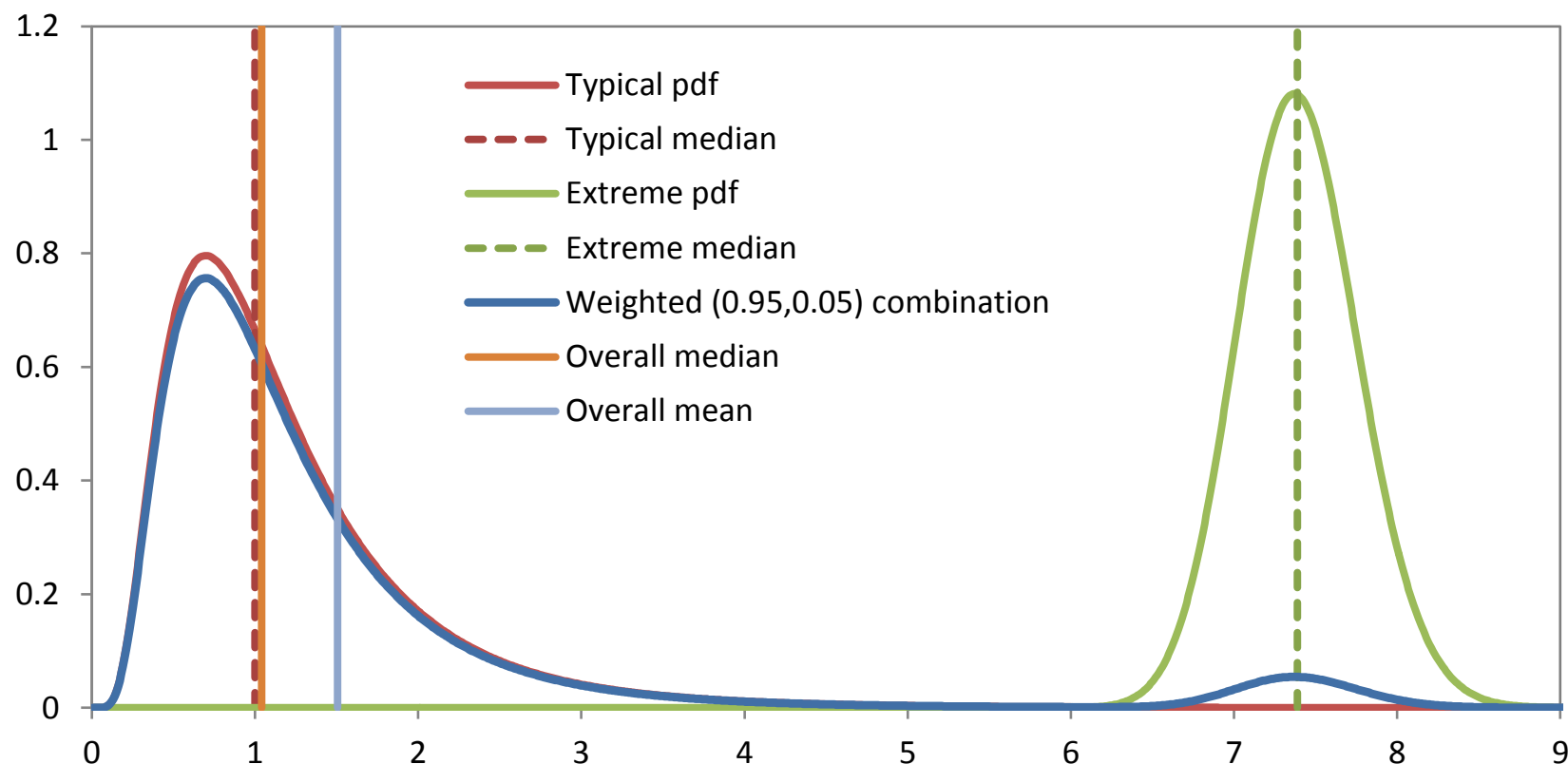
Analysis #1: other assumptions

- All fishing mortality was assumed to occur instantaneously at the start of the year, and was expressed in terms of a discrete annual exploitation rate U

- Natural mortality was expressed in terms of a discrete annual rate A

- A range of values (0.05, 0.10, 0.20, and 0.30) was considered for A

- Selectivity and maturity were assumed to be knife-edged, with the first age of full selectivity equal to the first age of full maturity

- Growth parameters were scaled so that $U_{35\%}=A$ and $B_{100\%}=1$

- Values of all parameters and variables *except the median and mean of the recruitment distribution* were assumed to be known without error

- The median and mean of the recruitment distribution were re-estimated in each year of the simulation, after a "burn-in" time of 30 years

  - i.e., no assessment was conducted during the first 30 years, so the first assessment has 30 years of recruitment data from which to estimate the median and mean

NOAA FISHERIES

Analysis #1: simulation procedure

- For each value of A and each estimator, 10,000 simulations were conducted
- Each simulation was initialized by assuming that the population was in equilibrium at 50% of the B_{MSY} proxy of $B_{35\%}$, where $B_{35\%}$ was scaled in terms of the respective estimator
  - i.e., $B_{35\%}=0.35 \times R_{med} \times SPR_{F=0}$ or $B_{35\%}=0.35 \times R_{ave} \times SPR_{F=0}$
- The maximum age in the population was defined as the age at which only 0.1% of a cohort would remain in an equilibrium unfished population (where cohort size is measured at the age of recruitment), and so was different for each value of A
- In each simulation, the population was projected forward for a number of years at least twice as great as the maximum age
- Alternative values for the Tier 3 reference points were computed for each of the two estimators (median and mean recruitment)
- Catch was assumed to equal maxABC under the Tier 3 control rule in all years

Analysis #1: performance metrics

- The following performance metrics were tabulated for each value of A and each of the two candidate estimators:
  - Short-term (first 10 years) and long-term (last 10% of the time series) means and standard deviations of:
    - Relative biomass (= biomass/$B_{40\%}$)
    - Relative exploitation (= exploitation/$U_{40\%}$)
    - Relative catch (=catch/$C_{40\%}$)
  - Five statistics pertaining to rebuilding time (to the B_{MSY} proxy, again scaled in terms of the respective estimator):
    - Upper and lower bounds of the 95% confidence interval
    - Median
    - Mean
    - Standard deviation

NOAA FISHERIES

Analysis #1: results

- See Tables F.1 and F.2
- Compared to use of mean recruitment, use of the median resulted in:
    - Less variability in long-term (but not short-term) fishing effort and catch
    - Higher long-term (but not short-term) average catch
    - Lower short- and long-term average biomass
    - Higher short- and long-term average exploitation rates
    - Shorter and less variable rebuilding times
        - Recall that both the initial biomass and the rebuilding target were scaled relative to the respective estimator
- Other performance metrics showed little difference between use of median and mean

NOAA FISHERIES

Analysis #2: an even simpler model

- Examination of an even simpler model may help to illustrate some differences between using the median and using the mean

- Suppose, instead of the model described in Analysis #1, that recruitment can take *only two values*, one low and one high

- If recruitment is normalized relative to the low value, the distribution is defined by two parameters:

  - the ratio of the high recruitment to the low recruitment ($ratio\_of\_values$)

  - the probability that the high value will occur ($proportion\_high$)

- The overall mean recruitment is then given as $overall\_mean$ = $(1 \times proportion\_high) + (proportion\_high \times ratio\_of\_values)$
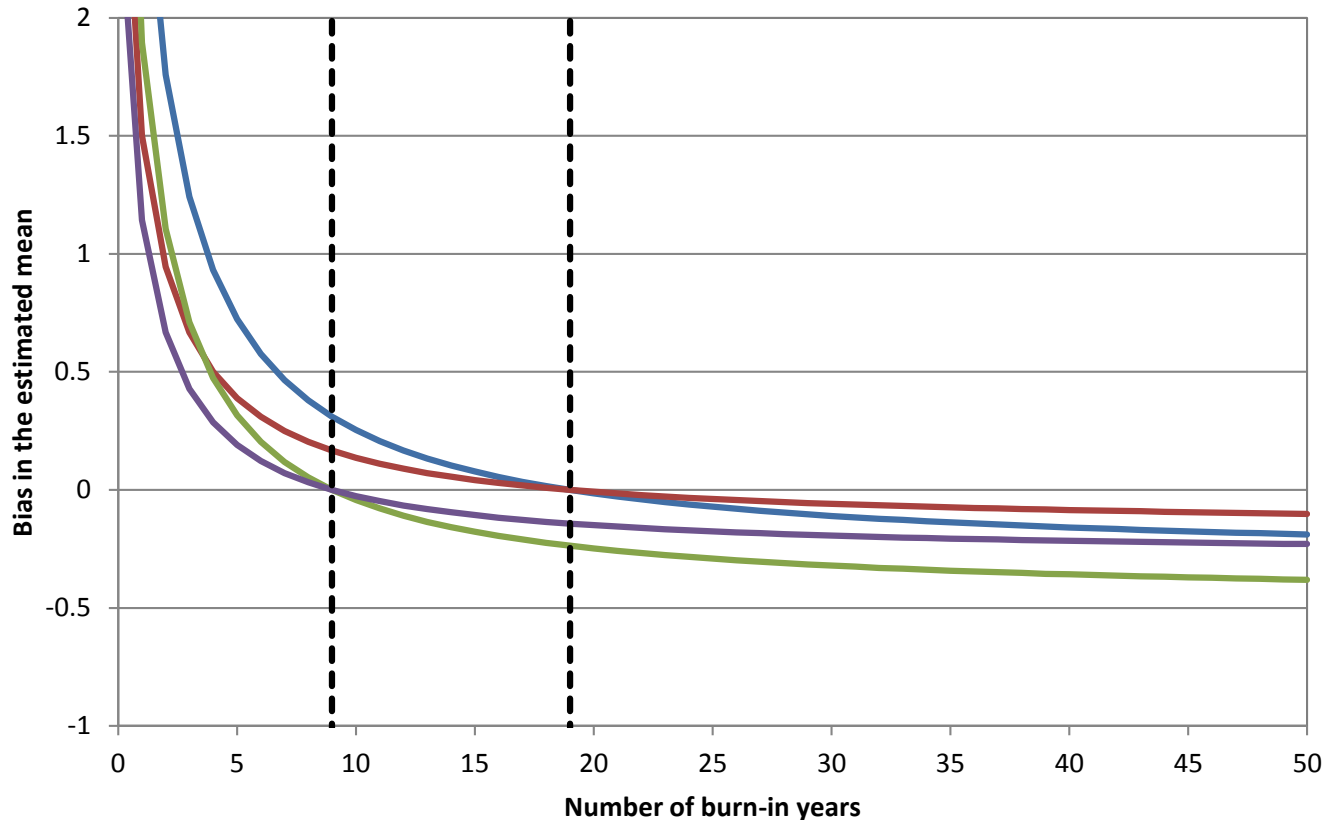
Analysis #2: worst-case scenarios (1 of 2)

- In this situation, the worst-case scenario in terms of bias in the estimated mean is for the stock to experience only low recruitments during the burn-in period of *nbur* years, followed by a high recruitment in the first year *after* the burn-in period:

$$bias_{mean} = \frac{nbur + ratio\_of\_values}{(nbur + 1) \cdot overall\_mean} - 1$$

- The worst of the worst-case scenarios occurs when there is no burn-in period at all, in which case the bias is positive and potentially large

- The bias then decreases continuously with *nbur* until it equals 0 at *nbur* = *proportion_high*[1] 1, after which it becomes negative and reaches an asymptote at *overall_mean*[1] 1

- The bias in the median is independent of *nbur*, being equal simply to the lower asymptote of $bias_{mean}$

- Some examples are shown on the next slide

Analysis #2: results (2 of 2)

- Blue: *ratio_of_values*=10, *proportion_high*=0.05; red: *ratio_of_values*=5, *proportion_high*=0.05; green: *ratio_of_values*=10, *proportion_high*=0.1; purple: *ratio_of_values*=5, *proportion_high*=0.1; vertical dash = no bias

U.S. Department of Commerce | National Oceanic and Atmospheric Administration | NOAA Fisheries | Page 22
This information is distributed solely for the purpose of pre-dissemination peer review under applicable information quality guidelines.
It has not been formally disseminated by the National Marine Fisheries Service and should not be construed to represent any agency determination or policy.

Conclusion

- This is a close call, but the Working Group recommends staying with the mean as the estimator of central tendency
- With only a couple (?) of exceptions, using the mean has not been a problem
  - But the exceptions could be significant
- Amending the FMPs would be a big deal
  - But maybe we will be required to do so anyway if the NS1 guidelines are revised
  - Revising SS would also be a big deal; revising Proj considerably less so
- Some performance gains w.r.t. catch are achieved by using the median; but, given the assumptions used in Analysis #1, so would anything that moves the control rule inflection point leftward

NOAA FISHERIES